

# A Steep Thermodynamical Selection Rule for Evolutionary Algorithms

Wei Qin Ying<sup>1</sup>, Yuanxiang Li<sup>1,2</sup>, Shujuan Peng<sup>2</sup>, and Weiwu Wang<sup>1</sup>

<sup>1</sup> State Key Lab. of Software Engineering, Wuhan University, Wuhan 430072, China

<sup>2</sup> School of Computer Science, Wuhan University, Wuhan 430079, China  
{weiqinying, yxli62}@yahoo.com.cn

**Abstract.** The genetic algorithm (GA) often suffers from the premature convergence because of the loss of population diversity at an early stage of searching. This paper proposes a steep thermodynamical evolutionary algorithm (STEA), which utilizes a steep thermodynamical selection (STS) rule. STEA simulates the competitive mechanism between energy and entropy in annealing to systematically resolve the conflicts between selective pressure and population diversity in GA. This paper also proves that the rule STS has the approximate steepest descent ability of the free energy. Experimental results show that STEA is both far more efficient and much stabler than the thermodynamical genetic algorithm (TDGA).

**Keywords:** Evolutionary algorithms, thermodynamics, selection rule, population diversity, free energy.

## 1 Introduction

The genetic algorithm (GA) is an optimization technique based on the mechanism of evolution by natural selection [1]. However, it has some disadvantages yet for solving large-scale combinatorial optimization problems because the astronomical size of search spaces with local optima often lead GA to extremely slow search and “premature convergence” [2]. The premature convergence also causes the low stability of GA. Whitley [3] argues that population diversity and selective pressure are the two primary factors in genetic search. Increasing selective pressure speeds up the search, while it also results in a faster loss of population diversity. Maintaining population diversity can help the search to escape local optima, while it offsets the effect of increasing selective pressure. These two factors are inversely related.

Some techniques on controlling population diversity have been proposed, such as scaling the fitness [1], sharing the fitness [1], and driving all individuals moving [4] etc. But they are not yet sufficiently systematical and effective for large-scale combinatorial problems. Kirkpatrick *et al.* [5] have proposed another general optimization algorithm called the simulated annealing (SA). SA controls search systematically by the cooling temperature and the Metropolis rule. Mori *et al.* [6,7] have proposed a method of combining SA and GA, called the thermodynamical genetic algorithm (TDGA). They introduce a greedy thermodynamical

selection rule in TDGA, getting a hint from the principle of minimal free energy. TDGA attempts to utilize the concepts of temperature and entropy in annealing to control population diversity systematically. TDGA is effective to some extent, while its performance is still inadequately stable and its computational cost is extremely high.

This paper proposes a steep thermodynamical evolutionary algorithm (STEA) to improve the stability and the computational efficiency of TDGA. STEA simulates the competitive mechanism between energy and entropy in annealing to systematically resolve the conflicts between selective pressure and population diversity in GA. The measurement of population diversity and the minimization process of free energy at each temperature are mended and a steep thermodynamical selection (STS) rule is proposed in STEA. The paper is organized as follows. We briefly review the thermodynamic background on STEA in Section 2, describe the outline of STEA in Section 3, and prove the approximate steepest descend ability of the rule STS in Section 4. Finally, the experimental results are presented in Section 5 and the conclusion is given in Section 6.

## 2 Brief Thermodynamic Background on STEA

In thermodynamics and statistical mechanics, annealing can be viewed as an adaptation process to optimize the stability of the final crystalline solid. In an annealing process, a metal, initially at high temperature and disordered, is slowly cooled so that the system at any temperature approximately reaches thermodynamic equilibrium [5]. As cooling proceeds, the system becomes more ordered and approaches a “frozen” ground state at the temperature  $T=0$ . There are a few observations about annealing which are helpful for STEA:

1. If the initial temperature is too low or cooling is done insufficiently slowly, the system may become quenched forming defects or trapped in a local minimum energy state.
2. Any change from non-equilibrium to equilibrium of the system at each temperature follows the principle of minimum free energy. In other words, the system will change spontaneously to achieve a lower total free energy and the system reaches equilibrium when its free energy seeks a minimum. In thermodynamics, the free energy  $F$  is defined as  $F = E - HT$ , where  $E$  is the energy of the system and  $H$  its entropy.
3. In thermodynamics, the entropy can quantificationally measure the energy dispersal of the particles in the system.
4. Any change of the system can be viewed as a result of the competition between its energy and its entropy. The temperature  $T$  determines their relative weights in the competition.

## 3 Steep Thermodynamical Evolutionary Algorithm

There are some deep and useful similarities between annealing in solids and convergence in GA. The population and the individuals in GA can be regarded

as the system and the particles. Then the mean of negative fitness, the population diversity and a controllable weight parameter can play the roles of the energy, the entropy and the temperature respectively. Every population state exactly full of global optima can be interpreted as the ground state. This analogy provides an approach for STEA to simulate the competitive mechanism between energy and entropy in annealing to systematically resolve the conflicts between selective pressure and population diversity in GA.

### 3.1 Measurement of Population Diversity

It is a critical part how to measure population diversity when introducing the competitive mechanism into GA. TDGA uses the sum of the information entropy at each locus, called a gene-based entropy in this paper. However, it has two disadvantages. Firstly, its repetitive calculations at all loci cause high computational costs. Secondly, the thermodynamic entropy can measure the energy dispersal of particles that is equivalent to the fitness dispersal of individuals in the population, while the gene-based entropy in TDGA can't measure the fitness dispersal. In this section, we propose a level-based entropy by grading the fitness.

**Definition 1.** Let  $S$  be the search space,  $f : S \rightarrow \mathbb{R}$  be the objective function, and  $X_r \in S$  be one individual. Then the individual energy  $e(X_r) = f(X_r)$  for minimum problems and  $e(X_r) = -f(X_r)$  for maximum problems. Assume  $e_u$  and  $e_l$  be respectively an upper bound and a lower bound of the individual energy. Then  $\pi = \{g_i | 0 \leq i \leq K - 1\}$  is called a level partition on  $[e_l, e_u]$  if

$$g_i = \left( \frac{2^{i-1} - 1}{2^{K-1} - 1} (e_u - e_l) + e_l, \frac{2^i - 1}{2^{K-1} - 1} (e_u - e_l) + e_l \right) \cap [e_l, e_u]. \tag{1}$$

We shall say that  $X_r$  is at level  $g_i$  if  $e(X_r) \in g_i$ .

**Definition 2.** Let  $P = (X_1, X_2, \dots, X_N) \in S^N$  be one population of size  $N$ ,  $\pi = \{g_i | 0 \leq i \leq K - 1\}$  be a level partition, and  $n_i$  be the number of individuals in  $g_i$  of population  $P$ . Then  $H(\pi, P)$  is called the level-based entropy of  $P$  for  $\pi$  where

$$H(\pi, P) = - \sum_{i=0}^{K-1} \frac{n_i}{N} \log_K \frac{n_i}{N}, \quad 0 \leq i \leq K - 1. \tag{2}$$

The value of the level-based entropy varies from 0 to 1 depending on the level distribution of  $P$ . It measures the fitness dispersal with very low computational costs.

### 3.2 Minimization of Free Energy at Each Temperature

**Definition 3.** For  $P \in S^N$ ,  $E(P)$  is called the population energy of  $P$  where

$$E(P) = \frac{1}{N} \sum_{r=1}^N e(X_r), \quad X_r \in P. \tag{3}$$

$F(\pi, T, P)$  is called the free energy of  $P$  at temperature  $T$  for partition  $\pi$  where

$$F(\pi, T, P) = E(P) - H(\pi, P)T. \tag{4}$$

The free energy is the driving force toward equilibrium in annealing. Similarly, we should force the population to minimize its free energy at each temperature  $T_k$  during evolution. However, there is only one generation of competition at each temperature  $T_k$  in TDGA. This insufficient competition only lowers the free energy very slightly, and the population can't reach the minimum free energy (or equilibrium) at  $T_k$ . In order to approach equilibrium, STEA holds  $L_k$  generations of competitions at  $T_k$ . These competitions at  $T_k$  form a *Markov* chain of length  $L_k$ , where  $L_k$  should grow with the hardness of problems.

### 3.3 Steep Thermodynamical Selection Rule

In order to minimize the free energy rapidly at each temperature  $T_k$ , we should design a thermodynamical selection rule to descend the free energy of the next generation most steeply. Its mission is to select  $N$  individuals from  $N$  parent individuals and  $M$  offspring individuals as the next generation with the minimum free energy. However, It is infeasible to exactly minimize the free energy for each generation because of the extremely high complexity  $O((N + K)C_{N+M}^N)$ . Hence, TDGA uses a greedy thermodynamical selection (GTS) rule with the complexity  $O(N^2K)$ . But its reliability can't be guaranteed. In this section, we propose a steep thermodynamical selection (STS) rule by assigning the free energy of the population to its individuals.

**Definition 4.** Let  $P = (X_1, X_2, \dots, X_N) \in S^N$  be one population of size  $N$  and  $\pi = \{g_i | 0 \leq i \leq K - 1\}$  be a level partition. For an individual  $X_r \in P$  at level  $g_d \in \pi$ , its free energy component in  $P$  at temperature  $T$  for  $\pi$  is defined as:

$$F_c(\pi, T, P, X_r) = e(X_r) + T \log_K \left( \frac{n_d}{N} \right), \tag{5}$$

where  $n_d$  is the number of individuals at  $g_d$  of  $P$ .

The steep selection rule **STS**( $\pi, T, P_t, O_t$ ) is described as follows:

1. Produce an interim population  $P'_t$  of size  $N + M$  by appending  $M$  individuals in the offspring population  $O_t$  to the parent population  $P_t$ .
2. Calculate the free energy component  $F_c(\pi, T, P'_t, X_r)$  for each individual  $X_r \in P'_t$ .
3. Pick the  $M$  individuals with the largest free energy components from  $P'_t$ .
4. Form the next generation  $P_{t+1}$  by removing these  $M$  individuals from  $P'_t$ .

STS has the lower complexity  $O((N + M)M)$ . It's proved in Section 4 that STS has the approximate steepest descent ability of the free energy.

### 3.4 Outline of STEA

Figure 1 provides the general outline of the whole STEA described above.

```

01 Create  $N$  individuals randomly as an initial population  $P_0$  and evaluate them;
02 Determine the energy bounds  $e_u$  and  $e_l$  for a level partition  $\pi$ ;
03 Configure the length  $L_k$  of the Markov chain at each temperature  $T_k$ ;
04  $T_0 = 10(e_u - e_l)$ ;
05  $t = 0$ ;  $k = 0$ ;
06 while(Termination_test( $P_t$ )==False)
07 {
08   for( $i = 0$ ;  $i < L_k$ ;  $i++$ )
09   {
10     Generate  $M$  offspring by uniform selection, crossover and mutation;
11     Organize these  $M$  offspring as an offspring population  $O_t$  and evaluate them;
12      $P_{t+1} = \mathbf{STS}(\pi, T_k, P_t, O_t)$ ;
13      $t = t + 1$ ;
14   }
15    $k = k + 1$ ;
16    $T_k = T_0/(1 + k)$ ;
17 }

```

Fig. 1. The outline of STEA

### 4 Approximate Steepest Descent Ability of STS

**Lemma 1.** Assume  $P_{t+1}^G$  is the next generation population with the exact minimum free energy and  $P_{t+1}$  is the next generation generated by STS. Then

$$\frac{1}{N} \sum_{X_r \in P_{t+1}} F_c(\pi, T_k, P'_t, X_r) - \frac{1}{N} \sum_{X_r \in P_{t+1}^G} F_c(\pi, T_k, P'_t, X_r) \leq 0, \tag{6}$$

$$0 \leq F(\pi, T_k, P_{t+1}) - \frac{1}{N} \sum_{X_r \in P_{t+1}} F_c(\pi, T_k, P'_t, X_r) \leq T_k \log_K ((N + M)/N), \tag{7}$$

$$0 \leq F(\pi, T_k, P_{t+1}^G) - \frac{1}{N} \sum_{X_r \in P_{t+1}^G} F_c(\pi, T_k, P'_t, X_r) \leq T_k \log_K ((N + M)/N). \tag{8}$$

Here we omit the proof of Lemma 1 for paper length limitation.

**Theorem 1.** Let  $y = \frac{M}{N+M}$  and  $D(\pi, T_k, P_t, O_t) = F(\pi, T_k, P_{t+1}) - F(\pi, T_k, P_{t+1}^G)$ . Then

$$\lim_{y \rightarrow 0} D(\pi, T_k, P_t, O_t) = 0. \tag{9}$$

*Proof.* Subtracting (8) from (7), we obtain an inequation:

$$D(\pi, T_k, P_t, O_t) \leq T_k \log_K ((N + M)/N) + \left( \frac{1}{N} \sum_{X_r \in P_{t+1}} F_c(\pi, T_k, P'_t, X_r) - \frac{1}{N} \sum_{X_r \in P_{t+1}^G} F_c(\pi, T_k, P'_t, X_r) \right). \tag{10}$$

Then we substitute (6) into (10) to get:

$$D(\pi, T_k, P_t, O_t) \leq T_k \log_K ((N + M)/N). \tag{11}$$

Since  $P_{t+1}^G$  has the exact minimum free energy, we also get:

$$D(\pi, T_k, P_t, O_t) \geq 0. \tag{12}$$

Furthermore, there exists such a limit:

$$\lim_{y \rightarrow 0} (T_k \log_K ((N + M)/N)) = 0. \tag{13}$$

Applying the *Squeeze Law* of limits, we can obtain (9) from (11), (12) and (13).

Theorem 1 asserts that the selection rule STS has the approximate steepest descent ability of the free energy if  $M \ll N$ . According to Theorem 1,  $M$  and  $N$  have been set to satisfy  $y < 0.1$  in all our experiments.

## 5 Experiments and Results

In this section, we present the experimental results on one instance of 0-1 knapsack problems, called KP2. KP2 is generated by the Pisinger’s algorithm [8] of constructing test instances with these parameters: *data rang*  $R=1000$ , *instance size*  $n=100$ , *problem type*  $t=$ *weakly correlated*, *instances sum*  $S=1000$  and *random seed*  $i=750$ . Note that here the profits  $p_j$ , the weights  $w_j$  and the capacity  $c$  are positive integers. We apply the simple genetic algorithm (SGA) with elitism, the steady state genetic algorithm (SSGA), TDGA and STEA to this instance. They all utilize the uniform crossover with its probability  $P_c = 0.8$ , the uniform mutation with its probability  $P_m = 0.05$ , and the same population size  $N = 80$ . The termination condition is satisfied when  $3.2 \times 10^6$  individuals are searched. The offspring population for SSGA, TDGA and STEA has the size  $M = 8$ . All experiments are performed on a Pentium-4 3.0G computer.

Assume that the items are ordered according to their efficiency such that  $p_i/w_i \geq p_j/w_j$  when  $i < j$ . Martello [9] has proved that the greedy value  $f_u = \sum_{j=1}^{b-1} p_j + [x_b p_b]$  is an upper bound of the objective function  $f$  where

$$\sum_{j=1}^{b-1} p_j \leq c, \sum_{j=1}^b p_j > c, \text{ and } x_b = (c - \sum_{j=1}^{b-1} w_j)/w_b. \tag{14}$$

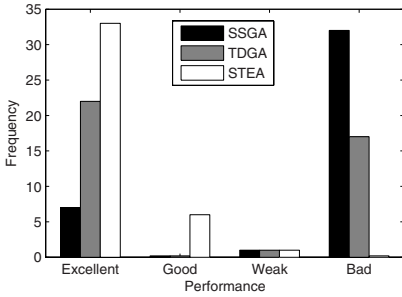
It’s also obvious that  $f_l = 0$  is an infimum of  $f$ . Therefor we can get an upper bound  $e_u = -f_l = 0$  and a lower bound  $e_l = -f_u = -38249$  of energy for the level partition  $\pi$  about KP2. The other parameters of STEA have the following values: level number  $K = \frac{1}{5}N = 16$  and chain length  $L_k = 10n = 1000$ .

There were respectively 40 trials on KP2 for each algorithm. Table 1 provides the statistic results of 40 trials for each algorithm, including the rate of hitting

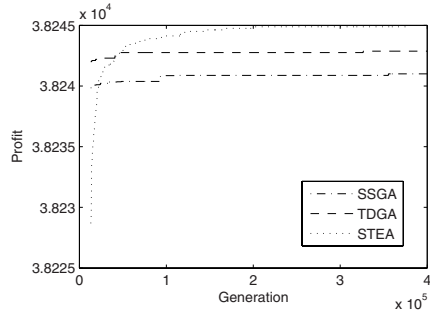
**Table 1.** Statistic results for four algorithms

Algorithm	Hitting rate	Worst	Mean	Best	Time	First hitting time
SGA	1/40	38145	38215.700	38245 <sup>†</sup>	152.2	>151.8
SSGA	8/40	38240	38241.000	38245 <sup>†</sup>	303.2	>254.0
TDGA	23/40	38240	38242.875	38245 <sup>†</sup>	56930.4	>26275.7
STEA	40/40	38245 <sup>†</sup>	38245.000 <sup>†</sup>	38245 <sup>†</sup>	191.8	=30.8

<sup>†</sup> optimal solution



**Fig. 2.** Performance distributions



**Fig. 3.** Convergence curves

optimum successfully, the worst result, the average result, the best result, average running time (seconds), and average time (seconds) of first hitting optimum. We score the performance during each trial for each algorithm according to the number  $s$  of searched individuals when first hitting optimum as follows: (1) *excellent* if  $s \leq 8 \times 10^5$ ; (2) *good* if  $8 \times 10^5 < s \leq 1.6 \times 10^6$ ; (3) *weak* if  $1.6 \times 10^6 < s \leq 3.2 \times 10^6$ ; (4) *bad* if  $s > 3.2 \times 10^6$ . Figure 2 shows the frequencies of four scores in 40 trials for SSGA, TDGA and STEA. Figure 3 shows their average convergence curves in 40 trials of the best individual of each generation. There are a few interesting observations which can be made on the basis of the experiments:

1. The results in Table 1 demonstrate clearly the stability of STEA. Its hitting rate is much higher than that of the other three algorithms. Note that for STEA the optimum was found in all 40 trials. Moreover, the quality of its solutions is averagely superior to that of the others due to its stability.
2. The above results also illustrate the high computational efficiency of STEA. STEA nearly spends the same running time as SGA. It contrasts clearly with the extremely high computational cost of TDGA, which is about 300 times as much as the others.
3. The first hitting time is a very significant performance goal for GA. It's very exciting that STEA has the far more rapid first hitting time than the others because of its high stability and low computational cost.
4. TDGA as well as SSGA presents the serious polarization phenomenon in Figure 2. It indicates that TDGA often get trapped in local optima.

However, STEA avoids this phenomenon due to keeping a systematical balance between selective pressure and population diversity successfully.

5. Figure 3 shows that the convergence speeds of SSGA and TDGA are faster than STEA at the early stage. However, they are in a nearly stagnant state after that stage and then the speed of STEA exceeds theirs. STEA gains the more rapid global convergence at the very little cost of the early stage.

## 6 Conclusions

This paper proposes a steep thermodynamical evolutionary algorithm (STEA), which utilizes a steep thermodynamical selection rule. STEA simulates the competitive mechanism between energy and entropy in annealing to systematically resolve the conflicts between selective pressure and population diversity in GA. The experimental results show that STEA not only speeds up the global convergence of TDGA remarkably at the very little cost of the early stage, but also improves the stability and the computational efficiency of TDGA greatly.

Further research will concentrate on the analysis of the convergence traits about STEA from the viewpoint of statistical mechanism.

**Acknowledgement.** This research is supported by the National Natural Science Foundation of China under Grant No.60473014.

## References

1. Goldberg, D.E.: Genetic algorithms in search, optimization, and machine learning. Addison-Wesley. (1989)
2. Su, X.H., Yang, B., Wang, Y.D.: A genetic algorithm based on evolutionary stable strategy. *Journal of Software*. **14**(11) (2003) 1863-1868 (in Chinese)
3. Whitley, D.: The GENITOR algorithm and selection pressure: Why rank-based allocation of reproductive trials is best. *Proc. of Int. Conf. on Genetic Algorithms*. (1989) 116-123
4. Li, Y.X., Zou, X.F., Kang, L.S., Michalewicz, Z.: A new dynamical evolutionary algorithm based on statistical mechanics. *J. Computer Science and Technology*. **18** (2003) 361-368
5. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by simulated annealing. *Science*. **220** (1983) 671-680
6. Mori, N., Yoshida, J., Kita, H., Nishikawa, Y.: A thermodynamical selection rule for the genetic algorithm. *IEEE Conf. on Evolutionary Computation*. (1995) 188-192
7. Kita, H., Mori, N., Nishikawa, Y.: Maintenance of Diversity by means of Thermodynamical Selection Rules for Genetic Problem Solving. *Proc. of Int. Symposium on Artificial Life and Robotics*. (1999) 330-333
8. Pisinger, D.: Core problems in knapsack algorithms. *Operations Research*. **47** (1999) 570-575
9. Martello, S., Toth, P.: Knapsack problems: algorithms and computer implementations. John Wiley & Sons. (1990)